Mathematics Department Stanford University
Math 171 Lecture Supplement
The Riemann and Lebesgue Integrals

Leon Simon

Autumn Quarter, 2012

These notes are meant as a quick introduction to the Riemann integral using step function terminology, followed by an almost-as-quick introduction to the Lebesgue integral, also via step functions. We include a proof of Lebesgue's theorem which precisely characterizes those functions which are Riemann integrable.

Contents

1 Preliminaries: Step Functions

Here we work in $\mathbb{R}^n$, where $n \geq 1$ is given. By an interval I in $\mathbb{R}^n$ we mean the cartesian product $I_1 \times I_2 \times \cdots \times I_n$, where each I_j is an interval in $\mathbb{R}$, thus each I_j is one of the following: $[a_j, b_j], (a_j, b_j), [a_j, b_j), (a_j, b_j]$, where $a_j \leq b_j$ are real. The volume $|I|$ of an interval $I = I_1 \times I_2 \times \cdots \times I_n$ is of course defined to be the product of the edge lengths. Thus $|I| = (b_1 - a_1) \times \cdots \times (b_n - a_n)$ in case I_j has end-points a_j, b_j with $a_j \leq b_j$, and the volume is zero if I is a degenerate interval (i.e. if $a_j = b_j$ for some j).

We say the interval I is closed if each factor I_j is a closed interval; thus a closed interval in $\mathbb{R}^n$ is one that can be written $[a_1, b_1] \times \cdots \times [a_n, b_n]$. Similarly, an open interval is one that can be written $I = (a_1, b_1) \times \cdots \times (a_n, b_n)$ ($= \emptyset$ if $a_j = b_j$ for some j).

From now on $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$ will be a fixed closed interval in $\mathbb{R}^n$ with $a_j < b_j$ for each $j = 1, \ldots, n$.

Definition: A *partition* $\mathcal{P}$ of R is the collection of closed intervals $I \subset R$ obtained by partitioning each of the edges of R; thus for each $j = 1, \ldots, n$ we select points $a_j = t_{j,0} < t_{j,1} < \cdots < t_{j,N_j} = b_j$ and then $\mathcal{P} = \{[t_{1,i_1-1}, t_{1,i_1}] \times [t_{2,i_2-1}, t_{2,i_2}] \times \cdots \times [t_{n,i_n-1}, t_{n,i_n}] : i_j \in \{1, \ldots, N_j\}$ for each $j = 1, \ldots, n\}$. The points $t_{j,0}, \ldots, t_{j,N_j}$ are called "the j-th edge points" of the partition $\mathcal{P}$. For any $I = [t_{1,i_1-1}, t_{1,i_1}] \times [t_{2,i_2-1}, t_{2,i_2}] \times \cdots \times [t_{n,i_n-1}, t_{n,i_n}] \in \mathcal{P}$ we let $\breve{I}$ denote the corresponding open interval $(t_{1,i_1-1}, t_{1,i_1}) \times (t_{2,i_2-1}, t_{2,i_2}) \times \cdots \times (t_{n,i_n-1}, t_{n,i_n})$, and $\partial I = I \setminus \breve{I}$.

Definition: φ is a *step function* on R if it is bounded and there is a partition $\mathcal{P}$ of R such that for each $I \in \mathcal{P}$ there is a real constant a_I such that $\varphi \equiv a_I$ on $\breve{I}$; thus

1.1 $$\varphi = \textstyle\sum_{I \in \mathcal{P}} a_I \chi_{\breve{I}} \text{ on } \cup_{I \in \mathcal{P}} \breve{I},$$

where we use the notation that, for any set $A \subset \mathbb{R}^n$, χ_A denotes the indicator function of A; thus $\chi_A(x) = 1$ if $x \in A$ and $\chi_A(x) = 0$ if $x \in \mathbb{R}^n \setminus A$.

Notice that in 1.1 we make no particular restriction on the values taken by φ on $\cup_{I\in\mathcal{P}}\partial I$ beyond the fact that $\varphi(x)$ should be a real number for each $x \in \cup_{I\in\mathcal{P}}\partial I$ and $\varphi|(\cup_{I\in\mathcal{P}}\partial I)$ should be bounded.

If φ and ψ are two such step functions on R, say with $\varphi = \sum_{I\in\mathcal{P}} a_I\chi_{\breve I}$ on $\cup_{I\in\mathcal{P}}\breve I$ and $\psi = \sum_{J\in\mathcal{Q}} b_J\chi_{\breve J}$ on $\cup_{J\in\mathcal{Q}}\breve J$, then the sum and difference is also a step function, as are $\max\{\varphi,\psi\}$, $\min\{\varphi,\psi\}$. One checks this by taking a common refinement $\mathcal{R}$ of the partitions $\mathcal{P},\mathcal{Q}$ (thus $\mathcal{R}$ is a partition of R with j-th edge points including all the j-th edge points of $\mathcal{P}$ and all the j-th edge points of $\mathcal{Q}$, $j = 1,\ldots,n$); then

1.2 $$\varphi = \textstyle\sum_{I\in\mathcal{R}} \widetilde a_I\chi_{\breve I} \text{ and } \psi = \sum_{I\in\mathcal{R}} \widetilde b_I\chi_{\breve I} \text{ on } \cup_{I\in\mathcal{R}}\breve I$$

for suitable $\widetilde a_I, \widetilde b_I$ (in fact $\widetilde a_J = a_I$ whenever $J \in \mathcal{R}$ with $J \subset I$ and $I \in \mathcal{P}$ and $\widetilde b_J = b_I$ whenever $J \in \mathcal{R}$ with $J \subset I$ and $I \in \mathcal{Q}$; this makes sense because every J in the common refinement $\mathcal{R}$ is contained in a unique $I \in \mathcal{P}$ and a unique $\widetilde I \in \mathcal{Q}$). Evidently it is also true that any scalar multiple of a step function is also a step function (indeed any product of step functions is a step function by 1.2) so in particular we have shown that the step functions form a subspace of the vector space of functions $f : R \to \mathbb{R}$. The vector space of step functions on R will be denoted

$$\mathcal{S}(R).$$

Naturally the integral of a step function φ as in 1.1 is defined by

1.3 $$\textstyle\int_R \varphi = \sum_{I\in P} a_I|I|,$$

where $|I|$ is the volume of I (i.e. the product of the edge lengths of I). Of course this definition is independent of the particular representation of step function used; thus if instead of 1.1 we use a different partition $\mathcal{Q}$ of R and if $\varphi = \sum_{J\in\mathcal{Q}} b_J\chi_{\breve J}$ on $\cup_{J\in\mathcal{Q}}\breve J$, then of course it turns out that $\sum_{I\in\mathcal{P}} a_I|I| = \sum_{J\in\mathcal{Q}} b_J|J|$; one can check this directly by using a common refinement $\mathcal{R}$ of $\mathcal{P},\mathcal{Q}$ as in 1.2 above, as follows: Since each $K \in \mathcal{R}$ is contained in a unique $I \in \mathcal{P}$ and $\varphi \equiv c_K$ on $\breve K$, where $c_K = a_I$, we have

$$\textstyle\sum_{K\in\mathcal{R}} c_K|K| = \sum_{I\in\mathcal{P}} a_I \sum_{K\in\mathcal{K} \text{ with } K\subset I}|K| = \sum_{I\in\mathcal{P}} a_I\,|I|,$$

and similarly

$$\textstyle\sum_{K\in\mathcal{R}} c_K|K| = \sum_{J\in\mathcal{Q}} b_J|J|,$$

so as claimed

$$\textstyle\sum_{I\in\mathcal{P}} a_I|I| = \sum_{J\in\mathcal{Q}} b_J|J|.$$

Using the formulae 1.2 it is evident that the integral so defined is linear on the step functions; that is, if φ,ψ are both step functions on R and if λ,μ are real constants then

1.4 $$\textstyle\int_R(\lambda\varphi + \mu\psi) = \lambda\int_R\varphi + \mu\int_R\psi, \quad \varphi,\psi \in \mathcal{S}(R).$$

It is also clear from the representation 1.2 that

1.5 $$\varphi,\psi \in \mathcal{S}(R) \text{ with } \phi \le \psi \Rightarrow \textstyle\int_R\phi \le \int_R\psi.$$

2 Riemann Integrable Functions

Given a bounded function $f : R \to \mathbb{R}$ we can define the *upper and lower Riemann integral of f*, denoted $\overline{\int}_R f$ and $\underline{\int}_R f$ respectively, by

$$\overline{\int}_R f = \inf_{\varphi \in \mathcal{S}(R), f \le \varphi} \int_R \varphi$$
$$\underline{\int}_R f = \sup_{\varphi \in \mathcal{S}(R), \varphi \le f} \int_R \varphi$$

and f is said to be Riemann integrable if it is bounded and $\overline{\int}_R f = \underline{\int}_R f$. Of course using 1.5 and the above definitions we have

$$\underline{\int}_R f \le \overline{\int}_R f$$

in any case, so to prove f is Riemann integrable it suffices to prove the reverse inequality.

We begin with a convenient criterion for checking Riemann integrability:

2.1 Lemma: *f is Riemann integrable $\iff$ for each given $\varepsilon > 0$ there are step functions φ, ψ with $\varphi \le f \le \psi$ on R and $\int_R \psi < \int_R \varphi + \varepsilon$ (i.e. $\int_R (\psi - \varphi) < \varepsilon$).*

Proof "⇐": We assume that for each $\varepsilon > 0$ there are step functions φ, ψ with $\varphi \le f \le \psi$ on R and $\int_R \psi < \int_R \varphi + \varepsilon$. Then $\underline{\int}_R f \le \overline{\int}_R f \le \int_R \psi \le \int_R \varphi + \varepsilon \le \underline{\int}_R f + \varepsilon$, whence $0 \le \overline{\int}_R f - \underline{\int}_R f < \varepsilon$, and hence (since this holds for each $\varepsilon > 0$) $\overline{\int}_R f = \underline{\int}_R f$, and so f is Riemann integrable as claimed.

Proof "⇒": For any $\varepsilon > 0$, by definition of upper and lower integral we can find step functions φ, ψ with $\varphi \le f \le \psi$ and with $\int_R \varphi > \underline{\int}_R f - \varepsilon/2$ and $\int_R \psi < \overline{\int}_R f + \varepsilon/2$. Thus if f is Riemann integrable (i.e. if the upper and lower integrals are equal), then we have $\int_R (\psi - \varphi) < \varepsilon/2 + \varepsilon/2 = \varepsilon$.

We now check a few general facts about the set of Riemann integrable functions.

2.2 Theorem: *Every continuous function $f : R \to \mathbb{R}$ is Riemann integrable.*

Proof: Let $\varepsilon > 0$. R is compact and a continuous function on a compact set is uniformly continuous, so there is $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ whenever $x, y \in R$ with $|x - y| < \delta$. So take a partition $\mathcal{P}$ of R such that each $I \in \mathcal{P}$ has diameter $< \delta$, and observe that then $\max_I f \le \min_I f + \varepsilon$ for each $I \in \mathcal{P}$. Define step functions φ, ψ on R as follows:

$$\varphi|\breve{I} = \min_I f, \ \psi|\breve{I} = \max_I f, \quad I \in \mathcal{P}$$

and take $\varphi = \min_R f$ and $\psi = \max_R f$ on the remaining points of R (i.e. on $\cup_{I \in \mathcal{P}} \partial I$). Then

$$\varphi \le f \le \psi \text{ and } \int_R \psi - \int_R \varphi = \textstyle\sum_{I \in \mathcal{P}} (\max_I f - \min_I f)|I| \le \varepsilon \sum_{I \in \mathcal{P}} |I| = \varepsilon |R|,$$

and hence f is Riemann integrable by Lemma 2.1.

2.3 Theorem: *The set of Riemann integrable functions on R is a vector space and the Riemann integral is a linear operator on that vector space (i.e. f, g Riemann integrable on R and $\lambda, \mu \in \mathbb{R} \Rightarrow \lambda f + \mu g$ is also Riemann integrable, and $\int_R (\lambda f + \mu g) = \lambda \int_R f + \mu \int_R g$).*

Proof: Let $\mathcal{R}(R)$ be the set of Riemann integrable functions on R. It suffices to prove

(1) $$f, g \in \mathcal{R}(R) \Rightarrow f + g \in \mathcal{R}(R) \text{ and } \textstyle\int_R (f+g) = \int_R f + \int_R g$$

and

(2) $$f \in \mathcal{R}(R) \text{ and } \lambda \in \mathbb{R} \Rightarrow \lambda f \in \mathcal{R}(R) \text{ and } \textstyle\int_R \lambda f = \lambda \int_R f.$$

To prove (1), let $\varepsilon > 0$ and, using Lemma 2.1, pick step functions $\varphi, \psi, \widetilde{\varphi}, \widetilde{\psi}$ such that $\varphi \le f \le \psi$, $\widetilde{\varphi} \le g \le \widetilde{\psi}$, and $\int_R(\psi - \varphi) < \varepsilon/2$, $\int_R(\widetilde{\psi} - \widetilde{\varphi}) < \varepsilon/2$, and so by virtue of the linearity 1.4 we have

$$\textstyle\int_R(\psi + \widetilde{\psi}) - \int_R(\varphi + \widetilde{\varphi}) < \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

and

$$\varphi + \widetilde{\varphi} \le f + g \le \psi + \widetilde{\psi},$$

so we deduce that $f+g$ is Riemann integrable by virtue of Lemma 2.1, and $\int_R f + \int_R g - \varepsilon \le \int_R(\psi + \widetilde{\psi}) - \varepsilon \le \int_R(\varphi + \widetilde{\varphi}) \le \int_R(f+g) \le \int_R(\psi + \widetilde{\psi}) \le \int_R(\varphi + \widetilde{\varphi}) + \varepsilon \le \int_R f + \int_R g + \varepsilon$, so

$$\textstyle\int_R f + \int_R g - \varepsilon \le \int_R(f+g) \le \int_R f + \int_R g + \varepsilon$$

for each $\varepsilon > 0$, whence $\int_R(f+g) = \int_R f + \int_R g$.
The proof of (2) is left as an easy exercise, again based on Lemma 2.1 and the linearity 1.4 (one has to break consideration into the cases $\lambda > 0, \lambda = 0, \lambda = -1$).

The Riemann integral also preserves ordering of functions as follows:

2.4 Theorem: *Suppose f, g are Riemann integrable on R and $f \le g$ (meaning $f(x) \le g(x)\, \forall x \in R$). Then $\int_R f \le \int_R g$.*

Proof: Let $\varepsilon > 0$. Since $\underline{\int}_R f = \int_R f$ we can find a step function φ with $\varphi \le f$ and $\int_R f \le \int_R \varphi + \varepsilon$, and since $\overline{\int}_R g = \int_R g$ we can find a step function ψ with $g \le \psi$ and $\int_R \psi \le \int_R g + \varepsilon$. Then $\varphi \le f \le g \le \psi$ and by 1.5 $\int_R f \le \int_R \varphi + \varepsilon \le \int_R \psi + \varepsilon \le \int_R g + \varepsilon + \varepsilon$; that is $\int_R f \le \int_R g + 2\varepsilon$, and since ε is arbitrary this implies $\int_R f \le \int_R g$.

3 Lebesgue measure zero

A key difference between Lebesgue theory and the theory of volume related to the Riemann integral is in the notion of measure zero. In the Riemann theory a set A is said to have volume zero (or "Jordan content zero") if for each $\varepsilon > 0$ there are open intervals $I_1, \ldots, I_N$ with $A \subset \cup_{j=1}^N I_j$ and $\sum_{j=1}^N |I_j| < \varepsilon$. In the Lebesgue theory, we allow infinitely many intervals:

3.1 Definition: A set $A \subset \mathbb{R}^n$ is said to have *Lebesgue measure zero* (or simply *measure zero*) if for each $\varepsilon > 0$ there are open intervals $I_1, I_2, \ldots$ such that $A \subset \cup_{j=1}^\infty I_j$ and $\sum_{j=1}^\infty |I_j| < \varepsilon$.

This makes a profound difference in the related theory; for example we have the following lemma:

3.2 Lemma. *Suppose $A_1, A_2, \ldots$ is a sequence of subsets of $\mathbb{R}^n$ such that each A_j has Lebesgue measure zero. Then $\cup_{j=1}^{\infty} A_j$ also has Lebesgue measure zero.*

Proof: Let $\varepsilon > 0$. For each j we can select open intervals $I_{j,1}, I_{j,2}, \ldots$ with $A_j \subset \cup_{k=1}^{\infty} I_{j,k}$ and $\sum_{k=1}^{\infty} |I_{j,k}| < 2^{-j}\varepsilon$. Then $\cup_{j=1}^{\infty} A_j \subset \cup_{j,k} I_{j,k}$ and $\sum_{j,k} |I_{j,k}| < \sum_{j=1}^{\infty} 2^{-j}\varepsilon = \varepsilon$

As a first indication of the importance of the notion of sets of measure zero, we have the following very elegant theorem of Lebesgue, which completely characterizes those functions which are Riemann integrable.

3.3 Theorem (Lebesgue's Theorem on the Riemann Integral.) *Let $f : R \to \mathbb{R}$ be a bounded function. f is Riemann integrable on $R \iff$ there is a set $A \subset R$ of Lebesgue measure zero such that $f : R \to \mathbb{R}$ is continuous at x for each point $x \in R \setminus A$.*

Caution: "$f : R \to \mathbb{R}$ is continuous at each $x \in R \setminus A$" is a much stronger condition than "$f|R \setminus A$ is a continuous function," and indeed $f|R \setminus A$ continuous is in general <u>not</u> sufficient to ensure that f is Riemann integrable even if A has measure zero. For example if we take $R = [0, 1]$, $A =$ the set of rationals in $[0, 1]$, then A has measure zero but the function f which is 1 on A and 0 on $R \setminus A$ is not Riemann integrable because evidently $\underline{\int}_R f = 0$ and $\overline{\int}_R f = 1$.

Proof of "$\Rightarrow$": Observe, by the definition of continuity, that f discontinuous at $y \in \breve{R} \iff \exists\, \varepsilon_0 > 0$ such that $\sup_I f - \inf_I f > \varepsilon_0\ \forall$ open interval I with $y \in I \subset R$, which is the same as saying there is a positive integer j such that $\sup_I f - \inf_I f > 1/j\ \forall$ open interval I with $y \in I \subset R$. Thus the set of discontinuities of $f|\breve{R}$ can be written $\cup_{j=1}^{\infty} S_j$, where

$$S_j = \{y \in \breve{R} : \sup_I f - \inf_I f > 1/j \text{ for every open interval } I \text{ with } y \in I \subset R\}.$$

By Lemma 3.2 we know that $\cup_{j=1}^{\infty} S_j$ has Lebesgue measure zero if each S_j has Lebesgue measure zero, so it is thus enough to prove that S_j has Lebesgue measure zero for each j.

Let $\varepsilon > 0$, $j \in \{1, 2, \ldots\}$. By Lemma 2.1 (with ε/j in place of ε), we can pick a partition $\mathcal{P}$ and corresponding $\varphi, \psi \in \mathcal{S}(R)$ with $\varphi \le f \le \psi$ and

$$\textstyle\sum_{I \in \mathcal{P}} (\sup_{\breve{I}} f - \inf_{\breve{I}} f)|I| \le \int_R (\psi - \varphi) < \varepsilon/j.$$

Since $\sup_{\breve{I}} f - \inf_{\breve{I}} f \ge 1/j$ whenever $S_j \cap \breve{I} \ne \emptyset$ (by definition of S_j), the above evidently implies $\sum_{\{I \in \mathcal{P} : S_j \cap \breve{I} \ne \emptyset\}} (1/j)|I| < \varepsilon/j$; that is,

$$(\ddagger) \qquad \textstyle\sum_{\{I \in \mathcal{P} : \breve{I} \cap S_j \ne \emptyset\}} |I| < \varepsilon.$$

But the intervals $\breve{I} \in \mathcal{P}$ with $\breve{I} \cap S_j \ne \emptyset$ cover all of S_j except for the set $E = \cup_{I \in \mathcal{P}} \partial I$ of measure zero, and hence $S_j \setminus E \subset \cup_{\{I \in \mathcal{P} : S_j \cap \breve{I} \ne \emptyset\}} \breve{I}$. Thus ($\ddagger$) proves that S_j can be covered by a finite union of intervals of total volume $< \varepsilon$ and hence S_j has Lebesgue measure zero as required.

Proof of "$\Leftarrow$": Let $\varepsilon > 0$ and cover the set of discontinuities of f by a union $\cup_{j=1}^{\infty} I_j$ of open intervals such that $\sum_{j=1}^{\infty} |I_j| < \varepsilon$. Then $K \equiv R \setminus \cup_{j=1}^{\infty} I_j$ is a compact set and f (as a function on R) is continuous at each point of this compact set. We can therefore assert that there is $\delta > 0$ such that

$$(*) \qquad |f(x) - f(y)| < \varepsilon \text{ whenever } x \in K,\ y \in R, \text{ and } |x - y| < \delta.$$

Notice that the statement is stronger than the standard fact that a continuous function on a compact set is uniformly continuous, because only the point x, and not necessarily the point y, is required to be in the compact set K—on the other hand, the proof using the Bolzano-Weierstrass theorem is almost identical to the usual Bolzano-Weierstrass proof of this standard fact, as follows: If there is $\varepsilon > 0$ such that $(*)$ fails for each $\delta > 0$ then it fails with $\delta = \frac{1}{k}$, $k = 1, 2, \ldots$, and hence there are points $x_k \in K, y_k \in R$ such that $|x_k - y_k| < \frac{1}{k}$ but $|f(x_k) - f(y_k)| \geq \varepsilon$. Then by the Bolzano-Weierstrass theorem we can find a convergent subsequence x_{k_j} with $x = \lim x_{k_j} \in K$. Since $|x_{k_j} - y_{k_j}| < \frac{1}{k_j} \leq \frac{1}{j}$ we also have $\lim y_{k_j} = x$, and so by continuity of f at x we have $f(x_{k_j}) - f(y_{k_j}) \to f(x) - f(x) = 0$, contradicting the fact that $|f(x_{k_j}) - f(y_{k_j})| \geq \varepsilon$ for each j.

Now, with δ as in $(*)$, we select any partition $\mathcal{P}$ of R with each edge of each $I \in \mathcal{P}$ having length $< \delta/\sqrt{n}$ (so diameter of each $I \in \mathcal{P}$ is $< \delta$). For any $I \in \mathcal{P}$ such that $I \cap K \neq \emptyset$ we have by $(*)$ that $\sup_I f - \inf_I f = \sup_{z_1, z_2 \in I}(f(z_1) - f(z_2)) = \sup_{z_1, z_2 \in I}((f(z_1) - f(y_I)) - (f(z_2) - f(y_I))) \leq \varepsilon + \varepsilon = 2\varepsilon$, where y_I is any point in $I \cap K$, while of course the sum of the volumes $|I|$ over the remaining $I \in \mathcal{P}$ is $\leq \varepsilon$ (because the remaining I have the property $I \cap K = \emptyset$ and hence all such $I \subset R \setminus K \subset \cup_j I_j$, and $\sum_j |I_j| < \varepsilon$). Thus we have

$$\textstyle\sum_{I \in \mathcal{P}} (\sup_I f - \inf_I f)|I| \leq 2\varepsilon|R| + (\sup_R f - \inf_R f)\varepsilon \leq 2\varepsilon(|R| + M), \quad M = \sup_R |f|.$$

Thus if we define $\varphi, \psi \in \mathcal{S}(R)$ by $\varphi|\breve{I} = \inf_I f$, $I \in \mathcal{P}$, and $\varphi = \inf_R f$ on $R \setminus \cup_{I \in \mathcal{P}} \breve{I}$, and $\psi|\breve{I} = \sup_I f$, $I \in \mathcal{P}$, and $\psi = \sup_R f$ on $R \setminus \cup_{I \in \mathcal{P}} \breve{I}$ then $\int_R (\psi - \varphi) \leq 2\varepsilon(|R| + M)$, so f is Riemann integrable by Lemma 2.1.

We conclude this section with two important lemmas about monotone sequences of step functions. Again the concept of sets of measure zero plays a key role. In the statement of these lemmas, and subsequently, we use the following terminology:

Terminology: We say that a property holds *"almost everywhere"* or *"for almost every point"* (abbreviated "a.e.") in a subset $\Omega \subset \mathbb{R}^n$ if there is a set $S \subset \Omega$ of Lebesgue measure zero such that the property in question holds at each point of $\Omega \setminus S$.

Thus for example $f_k(x) \to 0$ a.e. $x \in \Omega$ means that there is a set $S \subset \Omega$ with Lebesgue measure zero such that $\lim f_k(x) = 0$ for every $x \in \Omega \setminus S$.

The first lemma relates to decreasing sequences of non-negative step functions:

3.4 Lemma. *Suppose $\varphi_k \in \mathcal{S}(R)$ with $0 \leq \varphi_{k+1} \leq \varphi_k \,\forall\, k = 1, 2, \ldots$. Then $\int_R \varphi_k \to 0 \iff \varphi_k(x) \to 0$ a.e. $x \in R$.*

Proof of "⇒": Let $\alpha > 0$ and $S_\alpha = \{x \in R : \lim \varphi_k(x) \geq \alpha\}$. Then $S_\alpha \subset \{x \in R : \varphi_k(x) \geq \alpha\}$ for each k, so, with $\mathcal{P}_k$ any partition of R such that $\varphi_k|\breve{I}$ is constant for each $I \in \mathcal{P}_k$, we have $\alpha \sum_{I \in \mathcal{F}_k} |I| \leq \int_R \varphi_k$, where $\mathcal{F}_k = \{I \in \mathcal{P}_k : \varphi_k|\breve{I} \geq \alpha\}$, so $S_\alpha \subset (\cup_{I \in \mathcal{F}_k} \breve{I}) \cup (\cup_{\ell=1}^\infty (\cup_{I \in \mathcal{P}_\ell} \partial I))$ and $\sum_{I \in \mathcal{F}_k} |I| \leq \alpha^{-1} \int_R \varphi_k \to 0$. Since $\cup_{\ell=1}^\infty (\cup_{I \in \mathcal{P}_\ell} \partial I)$ has measure zero by Lemma 3.2, this shows that S_α has measure zero for each $\alpha > 0$. Then, by another application of Lemma 3.2, $\{x \in R : \lim \varphi_k(x) > 0\} = \cup_{j=1}^\infty S_{1/j}$ has measure zero as claimed.

Proof of "⇐": Let $\varepsilon > 0$. For each $k = 1, 2, \ldots$ let $\mathcal{P}_k$ be a partition of R such that φ_k is constant on $\breve{I}$ for each $I \in \mathcal{P}_k$, and let E be a set of Lebesgue measure zero such that $\varphi_k(x) \to 0$ for each

$x \in R \setminus E$. Since $E \cup (\cup_{k=1}^{\infty} \cup_{I \in \mathcal{P}_k} \partial I)$ has Lebesgue measure zero (by 3.2), we can select open intervals $I_1, I_2, \ldots$ such that $E \cup (\cup_{k=1}^{\infty} \cup_{I \in \mathcal{P}_k} \partial I) \subset \cup_{j=1}^{\infty} I_j$ and $\sum_{j=1}^{\infty} |I_j| < \varepsilon$.

Now $K = R \setminus (\cup_{j=1}^{\infty} I_j)$ is a compact set and by construction $\varphi_k(x) \to 0$ for each $x \in K$ and hence for each given $x \in K$ we can find k_x such that $\varphi_{k_x}(x) < \varepsilon$ and also (since $x \notin \cup_{I \in \mathcal{P}_{k_x}} \partial I$) there is $J_x \in \mathcal{P}_{k_x}$ with $x \in \breve{J}_x$ and $\varphi_{k_x} | \breve{J}_x \equiv \text{const.} < \varepsilon$. Trivially $K \subset \cup_{x \in K} \breve{J}_x$, so by compactness there is a finite set of points $x_1, \ldots, x_N \in K$ such that $K \subset \cup_{j=1}^{N} \breve{J}_{x_j}$. Then if $k \geq \max\{k_{x_1}, \ldots, k_{x_N}\}$ we have $\varphi_k < \varepsilon$ on $\cup_{j=1}^{N} \breve{J}_{x_j}$ because $\varphi_k(\xi) \leq \varphi_{k_{x_j}}(\xi) < \varepsilon$ for each $\xi \in \breve{J}_{x_j}$ and each $k \geq k_{x_j}$. So let $\mathcal{Q}$ be a partition of R with the property that, for $i = 1, \ldots, n$, the i-th edge points of the partition $\mathcal{Q}$ include all the i-th edge points of the all the intervals in the partitions $\mathcal{P}_{k_{x_1}}, \ldots, \mathcal{P}_{k_{x_N}}$. Then each $J \in \mathcal{Q}$ is either such that $\breve{J} \subset \breve{J}_{x_j}$ for some j or else $J \cap (\cup_{j=1}^{N} \breve{J}_{x_j}) = \emptyset$. Let ζ be a step function which is ε on the former intervals and $M = \sup_R \varphi_1$ on the latter (and keep in mind that each J of the latter kind is contained in $R \setminus (\cup_{j=1}^{N} \breve{J}_{x_j}) \subset \cup_{j=1}^{\infty} I_j)$. So we have $\varphi_k \leq \zeta$ on R for all $k \geq \max\{k_{x_1}, \ldots, k_{x_N}\}$, and hence $\int_R \varphi_k \leq \int_R \zeta \leq \varepsilon |R| + M \sum_{J \in \mathcal{Q} : J \cap (\cup_{j=1}^{N} \breve{J}_{x_j}) = \emptyset} |J| \leq \varepsilon |R| + M \sum_{j=1}^{\infty} |I_j| < (|R| + M)\varepsilon$ for all $k \geq \max\{k_{x_1}, \ldots, k_{x_N}\}$. Thus we have shown $\lim_{k \to \infty} \int_R \varphi_k = 0$ as required.

The second lemma concerns increasing sequences of step functions and is proved in a manner similar to the proof of the easy direction ("$\Rightarrow$") of Lemma 3.4:

3.5 Lemma. *Suppose $\varphi_k \in \mathcal{S}(R)$ with $\varphi_k \leq \varphi_{k+1}$ for each $k = 1, 2, \ldots$ and suppose also that $\{\int_R \varphi_k\}_{k=1,2,\ldots}$ is bounded. Then $\{\varphi_k(x)\}_{k=1,2,\ldots}$ is bounded for a.e. $x \in R$.*

Proof: We have to prove that S has measure zero, where

$$S = \{x \in R : \{\varphi_k\}_{k=1,2,\ldots} \text{ is unbounded}\}.$$

We are given $\{\int_R \varphi_k\}_{k=1,2,\ldots}$ is bounded, and it is increasing by 1.5, hence $\lim \int_R \varphi_k \in \mathbb{R}$ and we can select a subsequence $\{\varphi_{k_j}\}_{j=1,2,\ldots}$ such that $\int_R (\varphi_{k_{j+1}} - \varphi_{k_j}) < 2^{-2j}$ for each $j = 1, 2, \ldots$. For any $\alpha > 1$ observe that

$$S \subset \cup_{j=1}^{\infty} S_j, \text{ where } S_j = \{x \in R : \varphi_{k_{j+1}}(x) - \varphi_{k_j}(x) \geq 2^{-j}\alpha\},$$

because $x \notin S_j$ for each j implies $\varphi_{k_N}(x) - \varphi_{k_1}(x) = \sum_{j=1}^{N-1} (\varphi_{k_{j+1}}(x) - \varphi_{k_j}(x)) < \alpha \sum_{j=1}^{N} 2^{-j} < \alpha$ for each $N \geq 2$.

On the other hand for each j, with $\mathcal{P}_j$ a partition of R such that $(\varphi_{k_{j+1}} - \varphi_{k_j}) | \breve{I} = \text{constant}$ for each $I \in \mathcal{P}_j$, we have $2^{-j}\alpha \sum_{I \in \mathcal{F}_j} |I| \leq \int_R (\varphi_{k_{j+1}} - \varphi_{k_j}) < 2^{-2j}$, where $\mathcal{F}_j = \{I \in \mathcal{P}_j : (\varphi_{k_{j+1}} - \varphi_{k_j}) | \breve{I} \geq 2^{-j}\alpha\}$, hence $\alpha \sum_{j=1}^{\infty} \sum_{I \in \mathcal{F}_j} |I| < \sum_{j=1}^{\infty} 2^{-j} = 1$ and by construction $\{\breve{I} : I \in \mathcal{F}_j\}$ covers all of $S_j \setminus (\cup_{I \in \mathcal{P}_j} \partial I)$. Since $\cup_{j=1}^{\infty} \cup_{I \in \mathcal{P}_j} \partial I$ has measure zero by Lemma 3.2, we can select open intervals $I_1, I_2, \ldots$ such that $\cup_{j=1}^{\infty} \cup_{I \in \mathcal{P}_j} \partial I \subset \cup I_j$ and $\sum_{j=1}^{\infty} |I_j| < 1/\alpha$. Then $(\cup_{j=1}^{\infty} \{\breve{I} : I \in \mathcal{F}_j\}) \cup \{I_1, I_2, \ldots\}$ is a collection of open intervals covering $\cup_{j=1}^{\infty} S_j (\supset S)$ and the sum of their volumes $\leq 1/\alpha + 1/\alpha = 2/\alpha$. Since α is arbitrarily large this shows that S has measure zero as claimed.

4 Definition and Properties of the Lebesgue Integral

We let $\mathcal{S}_+(R)$ denote the set of non-negative step functions (i.e. $\mathcal{S}_+(R) = \{\varphi \in \mathcal{S}(R) : \varphi \geq 0\}$).

$\mathcal{L}_+(R)$ denotes the set of functions $f : R \to [0, \infty)$ with the property that there exists an increasing sequence of $\{\varphi_k\}_{k=1,2,\ldots} \subset \mathcal{S}_+(R)$ with $\{\int_R \varphi_k\}_{k=1,2,\ldots}$ bounded and $f(x) = \lim \varphi_k(x)$ a.e. $x \in R$. $\mathcal{L}_+(R)$ is closed under the operations of addition and multiplication by non-negative scalars and also under the operation of taking max and min of two functions in $\mathcal{L}_+(R)$; thus

4.1 $\qquad \lambda, \mu \geq 0,\ g, h \in \mathcal{L}_+(R) \Rightarrow \lambda g + \mu h \in \mathcal{L}_+$ and also $\max\{g, h\}, \min\{g, h\} \in \mathcal{L}_+(R)$,

because if φ_k, ψ_k are increasing sequences in $\mathcal{S}_+(R)$ converging a.e. to g, h respectively with $\int_R \varphi_k$ and $\int_R \psi_k$ bounded, then $\lambda\varphi_k + \mu\psi_k$, $\max\{\varphi_k, \psi_k\} (\leq \varphi_k + \psi_k)$ and $\min\{\varphi_k, \psi_k\}$ are increasing sequences in $\mathcal{S}_+(R)$ with bounded integrals converging a.e. to $\lambda g + \mu h$, $\max\{g, h\}$ and $\min\{g, h\}$ respectively.

However $\mathcal{L}_+(R)$ is not a vector space because it is not closed under the operation of multiplication by -1. On the other hand

$$\mathcal{L}^1(R) = \mathcal{L}_+(R) - \mathcal{L}_+(R) \ (= \{f : f = g - h \text{ with } g, h \in \mathcal{L}_+(R)\})$$

is a vector space since it is a non-empty subset of the vector space of real-valued functions on R and it is closed under addition and multiplication by scalars—it is closed under multiplication by non-negative scalars because $\mathcal{L}_+(R)$ is (by 4.1), and it is also trivially closed under multiplication by -1.

We are going to define the Lebesgue integral on the vector space $\mathcal{L}^1(R)$. In order to do this we first need to define it on $\mathcal{L}_+(R)$:

4.2 Definition: For $f \in \mathcal{L}_+(R)$ we define the Lebesgue integral $\int_R f$ of f by

$$\int_R f = \lim \int_R \varphi_k,$$

where $\{\varphi_k\}_{k=1,2,\ldots}$ is an increasing sequence in $\mathcal{S}_+(R)$ such that $\{\int_R \varphi_k\}_{k=1,2,\ldots}$ is bounded and $f(x) = \lim \varphi_k(x)$ a.e. $x \in R$ (as in the definition of $\mathcal{L}_+(R)$).

Notice that of course $\lim \int_R \varphi_k$ exists because $\{\int_R \varphi_k\}$ is bounded by assumption, and it is monotone increasing by 1.5. However to be sure the definition makes sense we have to show that $\lim \int_R \psi_k = \lim \int_R \varphi_k$ for any other increasing sequence $\{\psi_k\}_{k=1,2,\ldots} \subset \mathcal{S}_+(R)$ with $\{\int_R \psi_k\}_{k=1,2,\ldots}$ bounded and $f(x) = \lim \psi_k(x)$ a.e. $x \in R$. This is in fact an easy consequence of Lemma 3.4, as follows: For each fixed $k = 1, 2, \ldots$ we have that $\{(\varphi_k - \psi_\ell)_+\}_{\ell=1,2,\ldots}$ is a decreasing sequence of non-negative step functions which converges to zero for a.e. $x \in R$, and hence by Lemma 3.4 (the "$\Leftarrow$" direction of that lemma) we have $\lim_{\ell\to\infty} \int_R (\varphi_k - \psi_\ell)_+ = 0$ for each k, and hence $\int_R \varphi_k - \int_R \psi_\ell = \int_R (\varphi_k - \psi_\ell) \leq \int_R (\varphi_k - \psi_\ell)_+ \to 0$ as $\ell \to \infty$. Thus $\int_R \varphi_k \leq \lim \int_R \psi_\ell$ for each k and hence $\lim \int_R \varphi_k \leq \lim \int_R \psi_k$. The reverse inequality is proved by interchanging φ_k and ψ_k.

Thus the definition of the Lebesgue integral makes sense on $\mathcal{L}_+(R)$ and evidently has the additivity property that

4.3 $\qquad g, h \in \mathcal{L}_+(R)$ and $\lambda, \mu \geq 0 \Rightarrow \lambda g + \mu h \in \mathcal{L}_+(R)$ and $\int_R \big(\lambda f + \mu g\big) = \lambda \int_R g + \mu \int_R h$,

by 4.1, Definition 4.2, and the linearity 1.4 of the integral on step functions.

We also observe that (again by Definition 4.2)

4.4 $\qquad f : R \to [0, \infty),\ f = 0$ a.e. in $R \Rightarrow f \in \mathcal{L}_+(R)$ and $\int_R f = 0$,

because the zero sequence $0, 0, \ldots$ of step functions converges a.e. to f.

We can now immediately extend the definition of the Lebesgue integral to the vector space $\mathcal{L}^1(R)$ as follows:

4.5 Definition: If $f \in \mathcal{L}^1(R)$ then we define $\int_R f = \int_R g - \int_R h$, where $g, h \in \mathcal{L}_+(R)$ are such that $f = g - h$.

We of course have to check that this definition is independent of which $g, h \in \mathcal{L}_+(R)$ we choose to represent f. So suppose $g, h, \widetilde{g}, \widetilde{h} \in \mathcal{L}_+(R)$ and $g - h = \widetilde{g} - \widetilde{h}$. Then $g + \widetilde{h} = \widetilde{g} + h$ and by the additivity 4.3 we then have $\int_R g - \int_R h = \int_R \widetilde{g} - \int_R \widetilde{h}$.

Notice that by 4.3 and the Definition 4.5 we have the linearity

4.6 $$f, g \in \mathcal{L}^1(R), \lambda, \mu \in \mathbb{R} \Rightarrow \lambda f + \mu g \in \mathcal{L}^1(R) \text{ and } \int_R \big(\lambda f + \mu g\big) = \lambda \int_R f + \mu \int_R g.$$

We now check some of the basic further properties of the Lebesgue integral on $\mathcal{L}^1(R)$:

First we show that changing an $\mathcal{L}^1(R)$ function on a set of measure zero neither changes the fact that the function is in $\mathcal{L}^1(R)$ nor does it change the value of the integral:

4.7 $$f \in \mathcal{L}^1(R),\ \widetilde{f} : R \to \mathbb{R}, \text{ and } \widetilde{f} = f \text{ a.e. } \Rightarrow \widetilde{f} \in \mathcal{L}^1(R) \text{ and } \int_R f = \int_R \widetilde{f}.$$

This is clear because $\widetilde{f} = f + (\widetilde{f} - f) = f + (\widetilde{f} - f)_+ - (\widetilde{f} - f)_-$ and both $(\widetilde{f} - f)_\pm$ are in $\mathcal{L}_+(R)$ and have integral zero by 4.4.

Next observe that if $f \in \mathcal{L}^1(R)$ with $f \geq 0$ then by definition we have $g, h \in \mathcal{L}_+(R)$ with $f = g - h$ and $g \geq h$ (because $f \geq 0$), and hence, by Definition 4.2 and by 1.5, $\int_R g \geq \int_R h$, and Definition 4.5 then implies $\int_R f = \int_R g - \int_R h \geq 0$. Thus $f \in \mathcal{L}^1(R)$, $f \geq 0 \Rightarrow \int_R f \geq 0$. Since $f, g \in \mathcal{L}^1(R)$ with $f \geq g \Rightarrow f - g \geq 0$ we can use the linearity 4.6 to conclude

4.8 $$f, g \in \mathcal{L}^1(R) \text{ with } f \geq g \Rightarrow \int_R f \geq \int_R g.$$

Next we observe that $\mathcal{L}^1(R)$ is closed under the operation of taking absolute values:

4.9 $$f \in \mathcal{L}^1(R) \Rightarrow |f| \in \mathcal{L}^1(R) \text{ and } \left|\int_R f\right| \leq \int_R |f|.$$

To check this we choose $g, h \in \mathcal{L}_+(R)$ with $f = g - h$ and observe that then $|f| = \max\{g, h\} - \min\{g, h\}$ and also $\max\{g, h\}, \min\{g, h\} \in \mathcal{L}_+(R)$ (by 4.1), and so $|f| \in \mathcal{L}^1(R)$ and, by the Definition 4.5 and by 4.8, $\int_R f = \int_R g - \int_R h \leq \int_R \max\{g, h\} - \int_R \min\{g, h\} = \int_R |f|$, and then the proof of 4.9 is completed by applying the same argument to $-f$. In view of 4.9 we also have

4.10 $$f, g \in \mathcal{L}^1(R) \Rightarrow \max\{f, g\}, \min\{f, g\} \in \mathcal{L}^1(R)$$

because $\max\{f, g\} = \frac{1}{2}(f + g + |f - g|)$ and $\min\{f, g\} = \frac{1}{2}(f + g - |f - g|)$.

The next property seems intuitively obvious but the proof requires Lemma 3.4:

4.11 $$f \in \mathcal{L}^1(R),\ f \geq 0,\ \int_R f = 0 \Rightarrow f = 0 \text{ a.e.}$$

To check 4.11 choose $g, h \in \mathcal{L}_+(R)$ with $f = g - h$, and then observe that by definition of $\mathcal{L}_+(R)$ we have increasing sequences $\{\psi_\ell\}_{\ell=1,2,\ldots}$ and $\{\varphi_k\}_{k=1,2,\ldots} \subset \mathcal{S}_+(R)$ with $\psi_\ell \to g$ a.e. and $\varphi_k \to h$

a.e., and hence, with $\lambda = \sup \psi_\ell$, $\lim_{k\to\infty} \int_R (\lambda - (\psi_\ell - \varphi_k)_+) = \int_R (\lambda - (\psi_\ell - h)_+)$ by Definition 4.2. Hence $\lim_{k\to\infty} \int_R (\psi_\ell - \varphi_k)_+ = \int_R (\psi_\ell - h)_+ \le \int_R f$, where at the last step we used the hypothesis $f \ge 0$. Since $\int_R f = 0$ we can then apply Lemma 3.4 (the "$\Rightarrow$" direction of that lemma with $(\psi_\ell - \varphi_k)_+$ in place of φ_k) to give $(\psi_\ell - h)_+ = \lim(\psi_\ell - \varphi_k)_+ = 0$ a.e., and hence $\psi_\ell - h \le 0$ a.e., so by taking the pointwise limit with respect to ℓ we have $f = g - h \le 0$ a.e., and hence $f = 0$ a.e. as claimed.

We'll now check that the Lebesgue integral exists and agrees with the Riemann integral whenever the Riemann integral exists—thus the Lebesgue integral *entirely subsumes* the Riemann integral. At the same time we'll give an alternate characterization (supplementary to Theorem 3.3) of exactly which of the Lebesgue integrable functions are Riemann integrable.

Using Lemma 2.1 with $\varepsilon = 1/k$ we see that if $f : R \to \mathbb{R}$ is Riemann integrable then there are step functions ψ_k, φ_k with $\varphi_k \le f \le \psi_k$ and $\int_R \psi_k - 1/k \le \int_R \varphi_k \le RI(f) \le \int_R \psi_k \le \int_R \varphi_k + 1/k$ for $k = 1, 2, \ldots$, where $RI(f)$ is the Riemann integral of f, and hence $\int_R \Psi_k$ and $\int_R \Phi_k$ both converge to $RI(f)$, where $\Psi_k = \min\{\psi_1, \ldots, \psi_k\}$ and $\Phi_k = \max\{\varphi_1, \ldots, \varphi_k\}$. Thus we have $\Phi_k \le f \le \Psi_k$, Φ_k is a bounded increasing sequence of step functions, and Ψ_k is a bounded decreasing sequence of step functions. Also $\int_R (\Psi_k - \Phi_k) \le \int_R (\psi_k - \varphi_k) \to 0$, so by Lemma 3.4 (the "$\Rightarrow$" direction of that lemma) we have $\Psi_k(x) - \Phi_k(x) \to 0$ a.e. $x \in R$. Since $\Phi_k \le f \le \Psi_k$ for all x we thus have $\lim \Phi_k(x) = \lim \Psi_k(x) = f(x)$ a.e. $x \in R$. But then $f \in \widetilde{\mathcal{L}}_+(R) \cap \widetilde{\mathcal{L}}_-(R)$ where $\widetilde{\mathcal{L}}_+(R)$ denotes the set of functions $f : R \to \mathbb{R}$ such that there is an increasing sequence $\{\varphi_k\}_{k=1,2,\ldots} \subset \mathcal{S}(R)$ with $\varphi_k \le f$ on R and $f(x) = \lim \varphi_k(x)$ a.e. $x \in R$ and $\widetilde{\mathcal{L}}_-(R) = \{-f : f \in \widetilde{\mathcal{L}}_+(R)\}$. Also, observe that, with $\lambda = \sup_R |\Phi_1|$, $\Phi_k + \lambda \in \mathcal{S}_+(R)$, so by definition of the Lebesgue integral on $\mathcal{L}_+(R)$ we have $\int_R (\lambda + f) = \lim_{k\to\infty} \int_R (\lambda + \Phi_k)$, hence $\int_R f = \lim_{k\to\infty} \int_R \Phi_k$ which is also the Riemann integral of f by construction. Thus we have proved:

4.12 Lemma. *The set $\mathcal{R}(R)$ of Riemann integrable functions $f : R \to \mathbb{R}$ is a linear subspace of $\widetilde{\mathcal{L}}_+(R) \cap \widetilde{\mathcal{L}}_-(R)$ and the Riemann integral and the Lebesgue integral coincide on $\mathcal{R}(R)$.*

Actually there is also a partial converse: If $f \in \widetilde{\mathcal{L}}_+(R) \cap \widetilde{\mathcal{L}}_-(R)$ then we can find an increasing sequence φ_k and a decreasing sequence ψ_k of step functions with $f(x) = \lim \varphi_k(x) = \lim \psi_k(x)$ a.e. $x \in R$. For each k, let $\mathcal{P}_k$ be a partition of R such that both φ_k and ψ_k are constant on each $\breve{I}$ with $I \in \mathcal{P}_k$. Evidently we must then have $\varphi_k|\breve{I} \le \psi_k|\breve{I}$ for each $I \in \mathcal{P}_k$. Let $M > 0$ be such that $-M < \varphi_1$ and $\psi_1 < M$ everywhere on R. Then we can define $\widetilde{\varphi}_k(x) = \varphi_k(x)$ for $x \in \cup_{I\in\mathcal{P}_k} \breve{I}$ and $\widetilde{\varphi}_k(x) = -M$ on $\cup_{I\in\mathcal{P}_k} \partial I$, and similarly $\widetilde{\psi}_k(x) = \psi_k(x)$ for $x \in \cup_{I\in\mathcal{P}_k} \breve{I}$ and $\widetilde{\psi}_k(x) = M$ on $\cup_{I\in\mathcal{P}_k} \partial I$. Then $\widetilde{\varphi}_k(x) \le \widetilde{\psi}_k(x)$ at every point and they converge to $f(x)$ a.e. Indeed if E is a set of measure zero such that both $\varphi_k(x)$ and $\psi_k(x)$ converge to $f(x)$ at every point of $R \setminus E$, then by construction $\widetilde{\varphi}_k(x)$ and $\widetilde{\psi}_k(x)$ both converge to $f(x)$ for every $x \in R \setminus (E \cup (\cup_{k=1}^\infty \cup_{I\in\mathcal{P}_k} \partial I))$. However $\widetilde{\varphi}_k, \widetilde{\psi}_k$ are not monotone sequences. To remedy this we define $\Phi_k = \max\{\widetilde{\varphi}_1, \ldots, \widetilde{\varphi}_k\}$ and $\Psi_k = \min\{\widetilde{\psi}_1, \ldots, \widetilde{\psi}_k\}$. These are respectively increasing and decreasing sequences of step functions with $-M \le \Phi_k \le \Psi_k \le M$ everywhere on R and $\Phi_k(x), \Psi_k(x)$ converge to $f(x)$ at every point $x \in R \setminus (E \cup (\cup_{k=1}^\infty \cup_{I\in\mathcal{P}_k} \partial I))$. Thus if we let $\widetilde{f}(x) = \lim_{k\to\infty} \Phi_k(x)$ for <u>every</u> $x \in R$, then we have $\Phi_k \le \widetilde{f} \le \Psi_k$ for each k, $\widetilde{f} = f$ a.e., and $\int_R (\Psi_k - \Phi_k) \le \int_R (\psi_k - \varphi_k) \to 0$, so $\widetilde{f}$ is Riemann integrable by Lemma 2.1. Thus we have proved:

4.13 Lemma. *If $f \in \widetilde{\mathcal{L}}_+(R) \cap \widetilde{\mathcal{L}}_-(R)$ then there is a Riemann integrable $\widetilde{f}$ with $\widetilde{f} = f$ a.e.*

Remark: Note that by combining Lemma 4.12 and Lemma 4.13 we have proved that $f \in \widetilde{\mathcal{L}}_+(R) \cap \widetilde{\mathcal{L}}_-(R) \iff f$ is a.e. equal to a Riemann integrable function $\widetilde{f}$, and in that case the Riemann integral of $\widetilde{f}$ is equal to the Lebesgue integral of f (also equal to the Lebesgue integral of $\widetilde{f}$ by 4.7).

We conclude this section with a discussion of some important convergence theorems. First:

4.14 Theorem (Monotone Convergence Theorem.) *If $\{f_k\}$ is an increasing sequence in $\mathcal{L}^1(R)$ with $\{\int_R f_k\}_{k=1,2,\ldots}$ bounded, then $\exists\, f \in \mathcal{L}^1(R)$ such that $f_k(x) \to f(x)$ a.e. $x \in R$ and $\int_R f_k \to \int_R f$. Furthermore we can take $f \in \mathcal{L}_+(R)$ if $f_k \in \mathcal{L}_+(R)$ for each k.*

Proof: We consider first Case 1: $f_k \in \mathcal{L}_+(R)$ for each $k = 1, 2, \ldots$. In this case, by definition of $\mathcal{L}_+(R)$, for each $k = 1, 2, \ldots$ we can find an increasing sequence $\{\varphi_{k,j}\}_{j=1,2,\ldots} \subset \mathcal{S}_+(R)$ with $\varphi_{k,j}(x) \to f_k(x)$ a.e. $x \in R$ and $\int_R (f_k - \varphi_k) \le 2^{-k}$, where $\varphi_k = \varphi_{k,1}$. Then $\int_R (\varphi_{k,j} - \varphi_k) \le 2^{-k}$, $k, j = 1, 2, \ldots$, and hence

$$\int_R \sum_{k=1}^{j} (\varphi_{k,j} - \varphi_k) \le 1 \quad \forall\, j = 1, 2, \ldots,$$

and also $\{\sum_{k=1}^{j} (\varphi_{k,j} - \varphi_k)\}_{j=1,2,\ldots}$ is an increasing sequence in $\mathcal{S}_+(R)$. So by Lemma 3.5 we see that $\{\sum_{k=1}^{j} (\varphi_{k,j}(x) - \varphi_k(x))\}_{j=1,2,\ldots}$ is bounded a.e. $x \in R$. Thus a.e. $x \in R$ there is a fixed constant C_x such that $\sum_{k=1}^{j} (\varphi_{k,j}(x) - \varphi_k(x)) \le C_x$ for all j, and hence, since $\{\varphi_{k,i}\}_{i=1,2,\ldots}$ is increasing, we have, a.e. $x \in R$,

$$\sum_{k=1}^{\infty} (\varphi_{k,i}(x) - \varphi_k(x)) \le C_x \;\forall\, i = 1, 2, \ldots.$$

Since $\lim_{i\to\infty} \varphi_{k,i}(x) = f_k(x)$ a.e. $x \in R$, by taking $\lim_{i\to\infty}$ of the partial sums of the above series we conclude

$$\sum_{k=1}^{\infty} (f_k(x) - \varphi_k(x)) \le C_x \;\text{ a.e. } x \in R.$$

In particular $f_k(x) - \varphi_k(x) \to 0$ a.e. $x \in R$, and hence $\Phi_k = \max\{\varphi_1, \ldots, \varphi_k\}$ is an increasing sequence in $\mathcal{S}_+(R)$ with $f_k(x) - \Phi_k(x) \to 0$ a.e. $x \in R$, because $\varphi_k \le \Phi_k \le f_k$ for each k. The latter inequality implies $\int_R \Phi_k \le \int_R f_k$ for each k, and so $\{\int_R \Phi_k\}_{k=1,2,\ldots}$ is bounded and, by Lemma 3.5, $\{\Phi_k(x)\}$ is bounded a.e. $x \in R$ and we can define $f \in \mathcal{L}_+(R)$ by taking $f(x) = \lim \Phi_k(x)$ for x such that $\{\Phi_k(x)\}_{k=1,2,\ldots}$ is bounded and $f(x) = 0$ elsewhere. Then $\lim f_k(x) = \lim \Phi_k(x) = f(x)$ a.e. $x \in R$, and, by Definition 4.2, $\int_R f = \lim \int_R \Phi_k$. Since $\Phi_k \le f_k$ on R and $f_k \le f$ a.e. on R we then also have $\int_R f_k \to \int_R f$. This completes the proof in Case 1.

Case 2: The general case when $\{f_k\}$ is an increasing sequence in $\mathcal{L}^1(R)$ with $\{\int_R f_k\}_{k=1,2,\ldots}$ bounded. In this case set $f_0 \equiv f_1$ and observe that, for $k = 1, 2, \ldots$, $f_k - f_{k-1} \in \mathcal{L}^1(R)$ and hence there are $g_k, h_k \in \mathcal{L}_+(R)$ with $f_k - f_{k-1} = g_k - h_k$, and by Definition 4.2 we can select $\varphi_k \in \mathcal{S}_+(R)$ with $\varphi_k \le h_k$ and $\int_R (h_k - \varphi_k) < 2^{-k}$. Then $g_k - \varphi_k = (h_k - \varphi_k) + (f_k - f_{k-1}) \ge 0$ and also

$$\int_R (g_k - \varphi_k) \le 2^{-k} + \int_R (f_k - f_{k-1}).$$

Thus with

$$G_k = \sum_{j=1}^{k} (g_j - \varphi_j), \quad H_k = \sum_{j=1}^{k} (h_j - \varphi_j)$$

we see that $\{G_k\}, \{H_k\}$ are increasing sequences in $\mathcal{L}_+(R)$ with $\{\int_R G_k\}_{k=1,2,\ldots}$, $\{\int_R H_k\}_{k=1,2,\ldots}$ bounded, and $G_k - H_k = f_k - f_1$. By Case 1 above we have $G, H \in \mathcal{L}_+(R)$ such that $G_k \to G$ and $H_k \to H$ a.e. in R and $\int_R G_k \to \int_R G$, and $\int_R H_k \to \int_R H$. So with $f = G - H + f_1$ we have $f \in \mathcal{L}^1(R)$, $f_k \to f$ a.e. and $\int_R f_k \to \int_R f$. This completes the proof of the monotone convergence theorem.

We can now prove the important corollary that for any sequence $\{f_k\} \subset \mathcal{L}^1(R)$ with $\{\int_R |f_k|\}_{k=1,2,\ldots}$ bounded, pointwise convergence a.e. to some f implies $f \in \mathcal{L}^1(R)$:

4.15 Corollary. *Suppose $\{f_k\}_{k=1,2,\ldots} \subset \mathcal{L}^1(R)$ with $\{\int_R |f_k|\}_{k=1,2,\ldots}$ bounded, and suppose there is $f : R \to \mathbb{R}$ with $\lim f_k(x) = f(x)$ a.e. $x \in R$. Then $f \in \mathcal{L}^1(R)$.*

Proof: Since we can write $f_k = f_{k_+} - f_{k_-}$, where $f_{k_+} = \max\{f_k, 0\}$ and $f_{k_-} = \max\{-f_k, 0\}$, so that $f_{k_\pm} \in \mathcal{L}^1(R)$ by 4.10 and $f_{k_+}(x) \to f_+(x)$ and $f_{k_-}(x) \to f_-(x)$ a.e. $x \in R$, we see that it suffices to prove the corollary under the additional assumption that $f_k \geq 0$ for each k. So for each $k, \ell = 1, 2, \ldots$ we assume $f_k \geq 0$ and we let $F_{k,\ell} = -\min\{f_k, f_{k+1}, \ldots, f_{k+\ell}\}$ and observe that the sequence $\{F_{k,\ell}\}_{\ell=1,2,\ldots}$ is monotone increasing and bounded above by 0, and of course in $\mathcal{L}^1(R)$ by 4.10. Thus by the Monotone Convergence Theorem $\inf\{f_k, f_{k+1}, \ldots\} \in \mathcal{L}^1(R)$ because $\inf\{f_k(x), f_{k+1}(x), \ldots\} = -\lim_{\ell\to\infty} F_{k,\ell}(x)$ for each $x \in R$. But clearly $\inf\{f_k, f_{k+1}, \ldots\}$ is increasing and it is $\leq f_k$ on R for each $k = 1, 2, \ldots$. It also evidently has limit $f(x)$ a.e. $x \in R$, so we can apply the Monotone Convergence Theorem again to infer $f \in \mathcal{L}^1(R)$ and $\int_R f = \lim \int_R \inf\{f_k, f_{k+1}, \ldots\}$.

4.16 Remarks (1): The above does *not* show that $\int_R f = \lim_k \int_R f_k$ and indeed this is not true in general—obviously, because e.g. it may not even be true that the bounded real sequence $\int_R f_k$ has a limit. The above proof in fact only shows that, for non-negative $f_k \in \mathcal{L}^1(R)$ with $f_k \to f$ a.e. and $\int_R f_k$ bounded,

$$(*) \qquad \inf_{\ell\geq k} f_\ell \in \mathcal{L}^1(R) \text{ and } \int_R f = \lim_{k\to\infty} \int_R \inf_{\ell\geq k} f_\ell \left(\leq \lim_{k\to\infty} \inf_{\ell\geq k} \int_R f_\ell\right).$$

For example in the case $n = 1$ it is easy to construct a sequence f_k of non-negative step functions such that $\int_R f_k = 1$ for each k and $f_k(x) \to 0\ \forall x \in R$, thus showing that equality does not hold in general in the inequality $\int_R f \leq \lim_{k\to\infty} \inf_{\ell\geq k} \int_R f_\ell$.

(2): The above corollary makes it elementary to check that various algebraic operations leave $\mathcal{L}^1(R)$ invariant. For example

$$\begin{cases} f, g \in \mathcal{L}^1(R) \text{ and } g \text{ bounded } \Rightarrow fg \in \mathcal{L}^1(R) \\ f, g \in \mathcal{L}^1(R) \text{ and } g \geq 1 \ \Rightarrow f/g \in \mathcal{L}^1(R) \end{cases}$$

which are easily checked by using the definition $\mathcal{L}^1(R) = \mathcal{L}_+(R) - \mathcal{L}_+(R)$ in combination with the corollary.

As Remark 4.16(1) above makes clear, we cannot in general conclude $\int_R f = \lim \int_R f_k$ under the hypotheses of 4.15. However if a sequence $\{f_k\}_{k=1,2,\ldots} \subset \mathcal{L}^1(R)$ has the additional property that there is $F \in \mathcal{L}^1(R)$ such that $|f_k| \leq F\ \forall k$ then, still assuming $f_k \to f$ a.e., $(*)$ can be applied to both $F + f_k$ and $F - f_k$ in order to give $\inf_{\ell\geq k} f_\ell, \sup_{\ell\geq k} f_\ell \in \mathcal{L}^1(R)$ and

$$\int_R F + \int_R f = \int_R F + \lim_{k\to\infty} \int_R \inf_{\ell\geq k} f_\ell \leq \int_R F + \lim_{k\to\infty} \inf_{\ell\geq k} \int_R f_\ell,$$

hence $\int_R f \leq \lim_{k\to\infty} \inf_{\ell\geq k} \int_R f_\ell$, and

$$\int_R F - \int_R f = \int_R F - \lim_{k\to\infty} \int_R \sup_{\ell\geq k} f_\ell \leq \int_R F - \lim_{k\to\infty} \sup_{\ell\geq k} \int_R f_\ell,$$

hence $\lim_{k\to\infty} \sup_{\ell\geq k} \int_R f_\ell \leq \int_R f \leq \lim_{k\to\infty} \inf_{\ell\geq k} \int_R f_\ell$, and so $\lim_{k\to\infty} \int_R f_k$ exists and is equal to $\int_R f$. That is, we have proved:

4.17 Theorem ("Dominated Convergence Theorem"). *Suppose* $f_k \in \mathcal{L}^1(R)$, $f : R \to \mathbb{R}$ *with* $f_k(x) \to f(x)$ *a.e.* $x \in R$, *and suppose there is a function* $F \in \mathcal{L}^1(R)$ *such that* $|f_k| \le F\, \forall k$. *Then* $f \in \mathcal{L}^1(R)$ *and* $\lim_{k\to\infty} \int_R f_k = \int_R f$.

5 The Spaces $\mathcal{L}^1(R)$ and $\mathcal{L}^2(R)$

We conclude these notes with a discussion of some important further properties of the space $\mathcal{L}^1(R)$, and we also introduce the linear subspace $\mathcal{L}^2(R)$.

First, for $f \in \mathcal{L}^1(R)$ we define

$$\|f\|_1 = \int_R |f|,$$

and observe that $\|\,\|_1$ has the first two properties of a norm: $\|\lambda f\|_1 = |\lambda|\|f\|_1$ and $\|f+g\|_1 \le \|f\|_1 + \|g\|_1$, but in place of the usual positivity we have only $\|f\|_1 \ge 0$ for each $f \in \mathcal{L}^1(R)$, because $\|f\|_1 = 0$ evidently does not imply $f = 0$ but rather that $f = 0$ a.e. by 4.11. We therefore classify $\|\,\|_1$ as a "seminorm" rather than a norm, but we keep in mind that the positivity of $\|\,\|_1$ fails in very mild way, namely (by 4.11)

$$\|f\|_1 = 0 \Rightarrow f(x) = 0 \text{ a.e. } x \in R.$$

Here we establish the following completeness property of $\mathcal{L}^1(R)$:

5.1 Theorem. *Suppose* $\{f_k\}_{k=1,2,\ldots}$ *is a Cauchy sequence with respect to the seminorm* $\|\,\|_1$ *above; that is, for each* $\varepsilon > 0$ *there is* N *such that* $\int_R |f_k - f_\ell| < \varepsilon$ *whenever both* $k, \ell \ge N$. *Then there is* $f \in \mathcal{L}^1(R)$ *such that* $\int_R |f_k - f| \to 0$ *and such that, for some subsequence* $\{f_{k_j}\}_{j=1,2,\ldots}$ *of* $\{f_k\}_{k=1,2,\ldots}$, $f_{k_j}(x) \to f(x)$ *a.e.* $x \in R$.

Proof: We use the definition of Cauchy sequence with $\varepsilon = 2^{-j}, j = 1, 2, \ldots$, in order to select k_j such that $\int_R |f_\ell - f_{k_j}| < 2^{-j}$ for all $\ell \ge k_j$, and by selecting these k_j successively we can at the same time arrange that $k_{j+1} > k_j$ for all $j \ge 1$, so that we have in particular that $\{f_{k_j}\}_{j=1,2,\ldots}$ is a subsequence of $\{f_k\}_{k=1,2,\ldots}$ and

$$\|f_{k_{j+1}} - f_{k_j}\|_1 < 2^{-j} \tag{1}$$

for each j. In view of (1) we thus have that $\sum_{j=1}^{i} (f_{k_{j+1}} - f_{k_j})_+$, $i = 1, 2, \ldots$ is an increasing sequence in $\mathcal{L}^1(R)$ with integrals bounded above by the fixed constant 1, hence by the Monotone Convergence Theorem (4.14) we have $g = \sum_{j=1}^{\infty} (f_{k_{j+1}} - f_{k_j})_+ \in \mathcal{L}^1(R)$ and has integral equal to $\lim_{i\to\infty} \int_R \sum_{j=1}^{i} (f_{k_{j+1}} - f_{k_j})_+$. Likewise $h = \sum_{j=1}^{\infty} (f_{k_{j+1}} - f_{k_j})_- \in \mathcal{L}^1(R)$ and has integral equal to $\lim_{i\to\infty} \int_R \sum_{j=1}^{i} (f_{k_{j+1}} - f_{k_j})_-$.

Thus with $f = g - h + f_{k_1}$ we have, a.e. in R,

$$|f_{k_{i+1}} - f| = |\sum_{j=1}^{i} (f_{k_{j+1}} - f_{k_j}) - g + h| \le (g - \sum_{j=1}^{i} (f_{k_{j+1}} - f_{k_j})_+) + (h - \sum_{j=1}^{i} (f_{k_{j+1}} - f_{k_j})_-),$$

so $f_{k_i}(x) \to f(x)$ a.e. $x \in R$ and $\|f_{k_{i+1}} - f\|_1 \to 0$. Of course then $\lim_{k\to\infty} \|f - f_k\|_1 = 0$ because $\{f_k\}$ is Cauchy with respect to the $\mathcal{L}^1$ seminorm.

We conclude this discussion by introducing the space $\mathcal{L}^2(R)$:

5.2 Definition: $\mathcal{L}^2(R) = \{f \in \mathcal{L}^1(R) : f^2 \in \mathcal{L}^1(R)\}$.

At first sight it is not obvious that $\mathcal{L}^2(R)$ is a linear subspace of $\mathcal{L}^1(R)$, but one can very easily check this using the Dominated Convergence Theorem and Remark 4.16(2) as follows: By 4.16(2), $f, g \in \mathcal{L}^2(R) \Rightarrow f\,\frac{g}{1+k^{-1}|g|} \in \mathcal{L}^1(R)$ for each $k = 1, 2, \ldots$. Also, $|f\,\frac{g}{1+k^{-1}|g|}| \le |fg| \le \frac{1}{2}(f^2 + g^2) \in \mathcal{L}^1(R)$ and $f\,\frac{g}{1+k^{-1}|g|} \to fg$ pointwise, so by the Dominated Convergence Theorem $fg \in \mathcal{L}^1(R)$. That is

$$(\ddagger) \qquad f, g \in \mathcal{L}^2(R) \Rightarrow fg \in \mathcal{L}^1(R).$$

Since $(f+g)^2 = f^2 + g^2 + 2fg$ we thus have $f, g \in \mathcal{L}^2(R) \Rightarrow f + g \in \mathcal{L}^2(R)$. Since it is also clearly true that $f \in \mathcal{L}^2(R) \Rightarrow \lambda f \in \mathcal{L}^2(R)$, this completes the proof that $\mathcal{L}^2(R)$ is a linear subspace of $\mathcal{L}^1(R)$.

In view of (‡), we can define an "inner product" (or more correctly a "semi inner product") $\langle f, g\rangle$ on $\mathcal{L}^2(R)$ by

$$5.3 \qquad \langle f, g\rangle = \textstyle\int_R fg, \quad f, g \in \mathcal{L}^2(R).$$

Notice this has properties analogous to the dot product of vectors in $\mathbb{R}^n$:

(i) $\langle f, g\rangle = \langle g, f\rangle$ (symmetry)

(ii) $\langle \lambda f + \mu g, h\rangle = \lambda\langle f, h\rangle + \mu\langle g, h\rangle$ (linearity)

(iii) $\langle f, f\rangle \ge 0$ with equality $\iff f = 0$ a.e., (positivity)

where in checking (iii) "$\Rightarrow$" we used 4.11 (with f^2 in place of f). Using the usual argument (that, by (i)–(iii), $0 \le \langle f + tg, f + tg\rangle = t^2\langle g, g\rangle + 2t\langle f, g\rangle + \langle f, f\rangle$, which, for $\langle g, g\rangle \ne 0$, takes the non-negative minimum value of $(\langle f, f\rangle\langle g, g\rangle - \langle f, g\rangle^2)/\langle g, g\rangle$ when $t = -\langle f, g\rangle/\langle g, g\rangle$) we then have the Cauchy-Schwarz inequality:

$$5.4 \qquad \big|\langle f, g\rangle\big| \le \|f\|_2\|g\|_2,$$

where $\|f\|_2 = \sqrt{\langle f, f\rangle}$. $\|f\|_2$ is called "the inner product semi-norm" of $f \in \mathcal{L}^2(R)$. Since $\|f+g\|_2^2 = \|f\|_2^2 + \|g\|_2^2 + 2\langle f, g\rangle \le \|f\|_2^2 + \|g\|_2^2 + 2\|f\|_2\|g\|_2$ by 5.4, we thus have the triangle inequality

$$5.5 \qquad \|f + g\|_2 \le \|f\|_2 + \|g\|_2, \quad f, g \in \mathcal{L}^2(R).$$

We claim finally that $\mathcal{L}^2(R)$ is complete with respect to the semi-norm $\|\ \|_2$:

5.6 Theorem. *The space $\mathcal{L}^2(R)$ is complete relative to the inner product semi-norm; that is if $\{f_k\}_{k=1,2,\ldots} \subset \mathcal{L}^2(R)$ is Cauchy (i.e. for each $\varepsilon > 0$ there is N such that $k \ge \ell \ge N \Rightarrow \|f_k - f_\ell\|_2 < \varepsilon$) then there is $f \in \mathcal{L}^2(R)$ such that* $\lim \|f_k - f\|_2 = 0$.

Proof: Without loss of generality, because $f_k = f_{k+} - f_{k-}$ and $\{f_k\}$ Cauchy for $\|\ \|_2$ implies both $\{f_{k+}\}$ and $\{f_{k-}\}$ are Cauchy for $\|\ \|_2$, we can assume $f_k \ge 0$ on R for each k. Then observe $\|f_k^2 - f_\ell^2\|_1 = \int_R |f_k^2 - f_\ell^2| = \int_R |f_k - f_\ell|(f_k + f_\ell) \le \|f_k - f_\ell\|_2\|f_k + f_\ell\|_2 \le \|f_k - f_\ell\|_2(\|f_k\|_2 + \|f_\ell\|_2)$, where we used 5.4,5.5. Since $\|f_k\|_2, \|f_\ell\|_2 \le 1 + \|f_N\|$ for $k, \ell \ge N$, with N as in the definition of Cauchy sequence with $\varepsilon = 1$, we have thus shown that f_k^2 is a Cauchy sequence with respect to the semi-norm $\|\ \|_1$ of $\mathcal{L}^1(R)$ and so by Theorem 5.1 there is $h \in \mathcal{L}^1(R)$ with $\|f_k^2 - h\|_1 \to 0$ and, also by 5.1, there is a subsequence $f_{k_j}^2$ with $f_{k_j}^2(x) \to h(x)$ a.e. $x \in R$. Since $\|f_k - f_\ell\|_1 \le |R|^{1/2}\|f_k - f_\ell\|_2$ (by Cauchy-Schwarz again), we also have that f_k is Cauchy with respect to the semi-norm $\|\ \|_1$ of $\mathcal{L}^1(R)$, and hence there is an $f \in \mathcal{L}^1(R)$ such that $\|f_k - f\|_1 \to 0$. Theorem 5.1 also guarantees that there is a subsequence $\{f_{k_{j_\ell}}\}_{\ell=1,2,\ldots}$ of $\{f_{k_j}\}_{j=1,2,\ldots}$ which converges a.e. to f. But then $f_{k_{j_\ell}}^2$ converges a.e. to f^2 and so $h = f^2$ a.e., and we have proved that $f \in \mathcal{L}^2(R)$ and $\|f_k^2 - f^2\|_1 \to 0$, which implies $\|f_k - f\|_2^2 = \int_R |f_k - f|^2 \le \int_R |f_k - f|(f_k + f) = \int_R |f_k^2 - f^2| = \|f_k^2 - f^2\|_1 \to 0$.